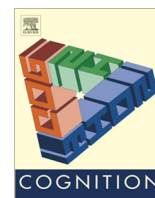


Contents lists available at [ScienceDirect](http://ScienceDirect.com)

Cognition

journal homepage: www.elsevier.com/locate/COGNITDeconstructing phonological tasks: The contribution of stimulus and response type to the prediction of early decoding skills[☆]Anna J. Cunningham^{a,1}, Caroline Witton^{a,b}, Joel B. Talcott^{a,b}, Adrian P. Burgess^{a,b}, Laura R. Shapiro^{a,*}^a School of Life and Health Sciences, Psychology, Aston University, UK^b School of Life and Health Sciences, Aston Brain Centre, Aston University, UK

ARTICLE INFO

Article history:

Received 6 August 2014

Revised 19 June 2015

Accepted 26 June 2015

Available online 10 July 2015

Keywords:

Phonological awareness

Decoding

Task

Reading

Measurement

Latent variable modelling

ABSTRACT

Phonological tasks are highly predictive of reading development but their complexity obscures the underlying mechanisms driving this association. There are three key components hypothesised to drive the relationship between phonological tasks and reading: (a) the linguistic nature of the stimuli, (b) the phonological complexity of the stimuli, and (c) the production of a verbal response. We isolated the contribution of the stimulus and response components separately through the creation of latent variables to represent specially designed tasks that were matched for procedure. These tasks were administered to 570 6 to 7-year-old children along with standardised tests of regular word and non-word reading. A structural equation model, where tasks were grouped according to stimulus, revealed that the linguistic nature and the phonological complexity of the stimulus predicted unique variance in decoding, over and above matched comparison tasks without these components. An alternative model, grouped according to response mode, showed that the production of a verbal response was a unique predictor of decoding beyond matched tasks without a verbal response. In summary, we found that multiple factors contributed to reading development, supporting multivariate models over those that prioritize single factors. More broadly, we demonstrate the value of combining matched task designs with latent variable modelling to deconstruct the components of complex tasks.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Although performance on tasks of phonological processing is strongly predictive of early reading (e.g. [Melby-Lervag, Lyster, & Hulme, 2012](#)), the underlying cognitive mechanisms that drive these relationships remain the subject of debate. One source of uncertainty is the complexity of phonological tasks, with many cognitive components potentially driving the associations with reading (as discussed by [Bowey, 2007](#); [Protopapas, 2014](#); [Ramus & Ahissar, 2012](#); [Snowling, Chiat, & Hulme, 1991](#)). Take, for example, three classic measures of phonological skill that have been found to be highly predictive of children's reading achievement;

phoneme isolation (e.g. 'what is the first sound in *san*?'; [Lervag, Braten, & Hulme, 2009](#); [Puolakanaho et al., 2007](#)), phoneme deletion (e.g. 'what is *san* without the *s*?'; [Hulme, Bowyer-Crane, Carroll, Duff, & Snowling, 2012](#); [Muter, Hulme, Snowling, & Stevenson, 2004](#)), and nonword repetition (e.g. 'say *san*'; [de Jong, 1998](#); [Nation & Hulme, 2011](#)). Each of these tasks share three key features: (a) the linguistic nature of the stimuli (which are usually words or pseudowords), (b) the phonological complexity of the stimuli (words can be segmented into phonemes), and (c) the response mode (which is nearly always by verbal report). In addition, there are demands on attention, short-term and working memory, and the ability to understand instructions during task execution. Since performance variability can result from the effect of any one, or combinations of these factors, it is typically not possible to discern the contributions of specific components. Of course, the issue of task complexity extends beyond reading research, and similar discussions have arisen in many areas of cognition (e.g., working memory, [Conway et al., 2005](#); executive function, [Hughes, 2011](#); language in relation to Theory of Mind, [Milligan, Astington, & Dack, 2007](#)).

The current study aimed to address these methodological and measurement issues by isolating the unique contribution of

[☆] This study was supported by the Economic and Social Research Council (ESRC, ES/H031685/1; PI Laura Shapiro). We are immensely grateful for the support and participation of the staff and pupils at the schools. Thanks go to the Research Assistants from Aston University who collected the data; <http://www.aston.ac.uk/alp/>.

* Corresponding author at: School of Life and Health Sciences, Psychology, Aston University, Aston Triangle, Birmingham B4 7ET, UK.

E-mail address: l.r.shapiro@aston.ac.uk (L.R. Shapiro).

¹ Present address: Centre for Research in Psychology, Behaviour and Achievement, Coventry University, Coventry CV1 5FB, UK.

stimulus and response mode respectively to the prediction of early decoding skills, while controlling as closely as possible for auxiliary task demands. This was achieved by using sets of carefully matched tasks as indicators for latent variables representing each component, and by partialing out memory and non-verbal reasoning ability.

1.1. Solutions to the problem of task complexity

A standard approach to disentangling complex tasks within cognitive psychology is to compare matched tasks with a common procedure. For example, [Vandermosten et al. \(2011\)](#) isolated the linguistic component of categorical perception tasks by comparing performance on a common task involving speech vs. non-speech stimuli matched for temporal cues (see [Banai & Ahissar, 2006](#); [Groth, Lachmann, Riecker, Muthmann, & Steinbrink, 2011](#) for similar paradigms). Another example comes from [Majerus, Linden, Mulder, Meulemans, and Peters \(2004\)](#), who isolated the role of sublexical knowledge on verbal short-term memory by comparing tasks using illegal vs. legal nonwords matched for procedure. These types of matched task designs are a fruitful way to examine the relative influences of different aspects of complex tasks. However, the studies relied on individual tasks to measure each construct and used relatively small samples. Outcomes may therefore be affected by heterogeneity in their samples, test sensitivity and method variance.

In contrast, latent variable modelling, or factor analysis (as exemplified in [Ramus, Marshall, Rosen, & van der Lely, 2013](#)) enables a more accurate estimate of an underlying skill by representing the commonalities among a range of measures, and extracting idiosyncratic task-specific factors as error variance. Additionally, structural equation modelling enables the correlations between latent variables to be explicitly modelled, providing an estimate of the unique contribution of each factor on an outcome ([Byrne, 2010](#); [Tabachnik & Fidell, 2007](#)). Many studies on the role of phonological skills in reading development have used latent variable approaches (see [Bowey, 2007](#) for a review). Nevertheless, the use of latent variables does not necessarily lead to a purer measure of each construct. Instead, the meaning of each latent variable is tied to the choice of indicators (measured variables) and the predictive power of each latent variable may depend on whether additional task demands are balanced equally between different constructs. In particular, the predictive power of a broad latent variable encompassing classic phonological tasks may be exacerbated by the breadth of skills they challenge (e.g. working memory, attention).

We propose that the role of specific components of complex tasks can be isolated using matched task designs in combination with latent variable modelling. For example, tasks can be created that follow a common procedure but vary in one crucial aspect (e.g., the response type: verbal or pointing using a touch screen). These closely matched tasks can then be used as indicators for correlated, but distinct latent variables (e.g., ‘verbal response’ and ‘non-verbal response’). The uniqueness of the component (in this case the verbal response) is extracted by the latent variable and then linked to an outcome (in this case decoding) using structural equation modelling. As the tasks vary along only one dimension, auxiliary demands are controlled for as closely as possible through the covariance between the factors (e.g., see [Kane et al., 2004](#) for a similar approach in the context of working memory).

1.2. The current study

The goal of the present study was to investigate the importance of three fundamental components of phonological tasks in the prediction of early decoding skills: The first two related to stimulus

(the linguistic nature and phonological complexity of the stimulus) and the third concerned response mode (verbal response).

A large sample of 6 to 7-year-old children (UK Year 2) was tested in order to capture an intermediate stage of reading development when phonological skills are most critical ([Ehri, 2005](#)). Decoding (regular and pseudoword reading) was used as the outcome measure as phonological processing more directly impacts on the reading proficiency of phonologically transparent items ([Snow & Juel, 2007](#)).

The classic tasks of phoneme isolation, deletion and nonword repetition were used as templates for four sets of novel tasks created through the systematic manipulation of stimulus type and response requirement; (1) tones with a non-verbal response (non-linguistic, non-verbal), (2) phonemes with a non-verbal response (linguistic, non-verbal), (3) phonemes with a verbal response (linguistic, verbal), and (4) pseudowords with a verbal response (phonologically complex, verbal). Tones with a verbal response were not included as pilot studies indicated that children could not reliably provide a verbal response to tonal stimuli. Similarly, pseudowords with a non-verbal response were not included as these tasks could not be matched in procedure to our other non-verbal response tasks. The consequences of using a design that was not full-factorial are explained in the discussion.

All twelve tasks were used as indicators for latent variables/factors that defined specific task components. We initially built a full model that combined both stimulus and response factors. However, it was not possible to calculate as extremely high correlations between factors (multicollinearity) caused by each task loading on both a stimulus and response factor meant that they could not be reliably separated in the prediction of decoding (see [Rigdon, 1995](#) for a discussion). Therefore, we tested two alternative models of our measures, structured either by stimulus or by response. The Stimulus model began with a tone factor (task-set 1), while the addition of a phoneme factor (task sets 2 and 3) represented the contribution of simple linguistic stimuli, and the addition of a pseudoword factor (task set 4) represented the contribution of complex linguistic stimuli to the prediction of decoding skills. Auxiliary task demands (such as understanding of instructions, attention and working memory load) were represented by the tone factor, leaving only stimulus-specific contributions to be made from the phoneme and pseudoword factors. The Response model began with a non-verbal response factor (task-sets 1 and 2), and the addition of the verbal response factor (task sets 3 and 4) represented the contribution of a verbal response. In both cases, the effects of verbal and visual-spatial short-term memory, and non-verbal reasoning were partialled out.

1.3. Predictions

All phonological theories of reading implicate the processing of phonological (linguistic) stimuli as central to the relationship with reading ([Ramus & Szenkovits, 2008](#); [Snowling & Hulme, 1994](#)). In addition, reading requires one to create and store accurate representations of speech units (words or pseudowords) comprising a series of segments ([Snowling, 2000](#); [Snowling & Hulme, 1994](#)). Therefore, we predict that both the linguistic nature and the phonological complexity of the stimuli in phonological tasks should drive the prediction of reading. In contrast, there is disagreement over whether giving a verbal response is critical. Research suggests that phonological tasks requiring a non-verbal response predict reading to a similar degree as those requiring a verbal response (e.g., [Gayan & Olson, 2003](#); [Hulslander et al., 2004](#)). However, the measures in these studies were not matched for the length of stimuli or processing demands, so a direct comparison of response type was not possible. However, other research has shown that paired associate learning tasks that required a

verbal output predicted word reading while equivalent tasks requiring a visual output (a drawing) contributed no additional variance (Litt, de Jong, van Bergen, & Nation, 2013). Therefore, based on this work, we predict that producing a verbal response should also be uniquely important for reading.

Our predictions for each model are that:

1.3.1. The Stimulus model

- (a) The simple linguistic nature of the stimulus: The factor (latent variable) representing phonemes will predict unique variance in decoding over and above the factor representing tones.
- (b) The phonological complexity of the stimulus: The factor representing pseudowords will predict unique variance in decoding over and above the factor representing phonemes.

1.3.2. The Response model

- (a) The factor representing a verbal response will predict unique variance in decoding over and above the factor representing a non-verbal response.

2. Method

2.1. Participants

All 585 children registered in Year 2 classes across 12 schools in the Birmingham area of the UK were invited to participate in the study. This is the third year of formal reading instruction in the UK. Parents were sent a letter informing them about the study and providing the opportunity to opt-out, while overall consent was given by the Headteacher. Twelve children did not participate, mainly due to parental opt-out. Of the remaining 573 children, complete data were available for 570 (48.5% boys) with a mean age of 6 years and 11 months (range = 6; 4–7; 6). Thirty children (5.2%) spoke English as an additional language but all had been in English-speaking education since the beginning of formal schooling. Most of the sample was White British (77%) with the remaining 23% coming from a variety of backgrounds including Indian, Pakistani and Black Caribbean. The proportion of children eligible for free school meals across the 12 schools ranged from 4.3% to 40.1%, with a mean of 22.32% (SD = 16.42) which is slightly higher than the national average of 18.3% (Department for Education, 2013).²

2.2. Exogenous (predictor) measures

All children were administered the Component Phonological Skills Assessment Scales (CPSAS) (Cunningham, Witton, Talcott, Burgess, & Shapiro, 2011). The CPSAS consisted of four task-sets that systematically varied the properties of the stimuli (tones, phonemes and pseudowords) and response type (non-verbal and verbal). For each task-set, there were three individual tasks (ISOLATION, REPETITION and DELETION) giving 12 scales in total (see Fig. 1).

- (1) Tone sequences with a non-verbal response (TonesNVR).
- (2) Phoneme sequences with a non-verbal response (PhonemesNVR).
- (3) Phoneme sequences with a verbal response (PhonemesVR).
- (4) Pseudowords with a verbal response (PseudowordsVR).

2.2.1. Task instructions

Each task consisted of 21 items, divided into 3 levels of increasing difficulty.

Level 1. Three practice items followed by 9 trials consisting of sequences of two stimuli (tones or phonemes), or individual single-syllable pseudowords. Children were required to isolate or delete the first stimulus of the tone or phoneme sequence (or initial phoneme of the pseudoword), repeat both stimuli of the tone or phoneme sequence in order (or repeat back the whole pseudoword).

Level 2. Two practice trials followed by 6 items consisting of sequences of three stimuli (tones or phonemes), or longer single-syllable pseudowords. Children were required to isolate or delete the last stimulus of the tone or phoneme sequence (or final phoneme of the pseudoword), repeat all three stimuli of the tone or phoneme sequence in order (or repeat back the whole pseudoword).

Level 3. The final level consisted of 2 practice items, followed by 6 items, consisting of sequences of 4 stimuli (tones and phonemes), or multi-syllable pseudowords. Requirements to isolate, delete or repeat were the same as for level 2.

2.2.2. Training associations

For the TonesNVR and PhonemesNVR tasks, children gave their responses by pressing one of 3 identical aliens presented horizontally on a screen. Children were given training at the start of each task-set so they could learn to associate each alien with its corresponding tone or phoneme. During training, the computer played one of the three tones or phonemes (each was played five times in random order) and the child had to press the alien that made that same sound. If they got 10 or more out of 15 correct, they passed the training and proceeded to the main tasks. If they did not pass the training, it was repeated up to three times, after which, the tasks were administered even if the last training session had been failed. 526 (92.3%) passed the TonesNVR training and 564 children (98.9%) passed the PhonemesNVR training.³

The PhonemesVR tasks were the same as PhonemesNVR tasks, except that the computer screen was facing the experimenter and the children provided their responses verbally. There was no training phase but all children were able to repeat each phoneme clearly at least once before the tasks began.

2.2.3. Stimuli

For the Tone tasks, the three stimuli were 300-ms pure tones of 500 Hz, 1000 Hz and 2000 Hz, produced with a 44.1 kHz sample rate and 16-bit encoding and presented with an inter-stimulus interval of 300 ms. Each tone was gated with a 10-ms raised cosine rise and fall time to prevent audible clicks at onset and offset. These frequencies were chosen because they are easily discriminated based on pitch, and because they all fall within a range dominated by the same underpinning neural mechanisms for perception (Moore, 2003). The 500 Hz tone was associated with the left-hand alien, the 1000 Hz tone with the central alien, and the 2000 Hz tone with the right-hand alien, i.e. in left-right order of increasing pitch. The order in which the recordings were played was the same across isolation, repetition and deletion.

For the PhonemesNVR and PhonemeVR tasks the three stimuli were the stop consonants /g/, /k/, /p/ presented for 500 ms, produced with a 44.1 kHz sample rate and 16-bit encoding, presented with an inter-stimulus interval of 300 ms. The /g/ was associated with the left-hand alien, the /k/ with the central alien, and the

² Multiple group models that split the schools into 2 groups (high SES; Free school meals = 1–10% – 4 schools) and (low SES; Free school meals = 31–40% – 3 schools) showed that when parameters were constrained to be equal between models, the fit was not significantly worse. Therefore, we can reasonably assume that the pattern of results is consistent across SES.

³ Two separate models comparing children who passed/did not pass the training showed that parameters were not significantly different between models.

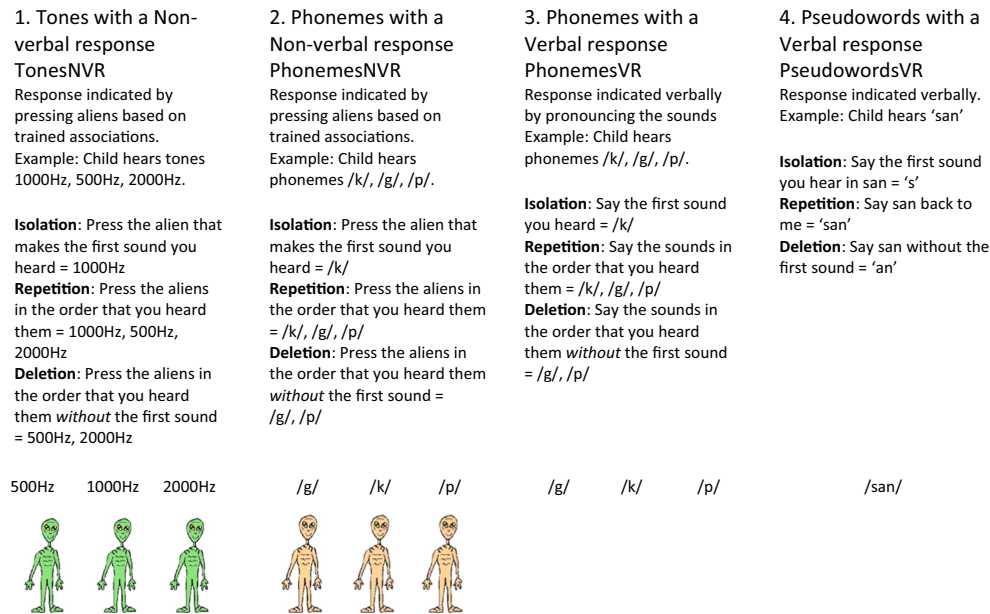


Fig. 1. CPSAS tasks. Screen-shots are provided for the non-verbal response tasks. Note that the letters and Hz shown above the aliens did not appear on the screen. Where children provided a verbal response, there was no visual display.

/p/ with the right-hand alien. The order of stimuli for each item (in terms of position of the aliens) was identical to the tone tasks. Stop consonants were chosen as these are the earliest acquired of the consonants (Kilminster & Laird, 1978). Phonemes were played from recordings made in Audacity (version 1.3.4-beta, 2011), selected at random from a pool of 28 recordings of two female voices.

For the PseudowordsVR tasks a selection of 21 pseudowords were used to cover a range of difficulty: 12 single-syllable pseudowords were taken from the YARC sound isolation task (Snowling et al., 2009) and 9 multi-syllable pseudowords (2–5 syllables) were taken from the Children's Test of Non-word Repetition (Gathercole, Willis, Baddeley, & Emslie, 1994). Pseudowords were recorded in Audacity (2011) with a 44.1 kHz sample rate and 16-bit encoding.

All tasks were administered on a notebook computer. The tone and phoneme tasks were developed using the 'pygame' module in Python to create interactive touch-screen displays (Sweigart, 2010). The pseudoword tasks were programmed in E-prime (Schneider, Eschman, & Zuccolotto, 2002). Sample-specific reliabilities are reported in Table 1.

2.3. Endogenous (outcome) measures

2.3.1. Decoding

Three measures of decoding were used.

Regular word reading from the British Ability Scales-2 school-age battery (Elliot, Smith, & McCulloch, 1996). Children were asked to read as many words as possible from a list of increasing difficulty containing a mixture of regular (43) and irregular words (47). Regular words were defined as those consistent with letter-sound mappings taught to the children in the sample (Department for Education and skills, 2007). The total for regular words only was used for the models.

Regular word reading from the Diagnostic Test of Word Reading Processes (Forum for Research in Literacy and Language, 2012) consisting of 30 regular words was administered.

Nonword reading from the Diagnostic Test of Word Reading Processes (Forum for Research in Literacy and Language, 2012) consisting of 30 nonwords was administered.

2.4. Control measures

Non-verbal reasoning was assessed using the British Ability Scales 2 Matrices test.

Verbal short-term memory was assessed using the British Ability Scales 2 Recall of Digits sub-test (Elliot et al., 1996).

Visuo-spatial short-term memory was assessed using a computerised version of Corsi blocks (the touchscreen programme from De Lillo, 2004 was adapted to follow the standard procedure in; Kessels, van Zandvoort, Postma, Kappelle, & de Haan, 2000).

2.5. Procedure

All children were tested during the second trimester of Year 2 (US 2nd Grade) between January and March. Tasks were administered by a trained team of 12 research assistants in a quiet area outside the classroom. Testing conditions and experimenter consistency were monitored twice per research assistant by the first author. The four task-sets were administered in fixed order from easiest to most difficult over 4 sessions lasting approximately 20 min each (PhonemesVR, PseudowordsVR, PhonemesNVR, and TonesNVR). Tasks were administered through headphones (Sennheiser, HD 25-111) at a comfortable hearing level that was equated across computers.

2.6. Data preparation

Most variables were approximately normally distributed with skew and kurtosis values below 1 (see Table 1). Large negative skew and positive kurtosis values were evident for the verbal and non-verbal phoneme isolation tasks. Inverse transformation ($1/(X_{\text{Highest}} - X_i)$) of these variables notably normalised the distributions so the transformed scores were used in the analyses. The resulting dataset displayed multivariate normality (critical ratio of multivariate kurtosis = 1.68). Finally, the data were screened for multivariate outliers. One Mahalanobis D^2 value stood

Table 1
Descriptive statistics for the predictor, control, and outcome variables.

Variable		Number of items correct		Skew	Kurtosis	Cronbach's alpha
Task-set	Task	Max.	M (SD)			
TonesNVR	Isolation	21	12.5 (5.2)	-.65	-.43	.86
TonesNVR	Repetition	21	5.6 (3.8)	.73	.16	.81
TonesNVR	Deletion	21	6.2 (3.4)	.94	.94	.75
PhonemesNVR	Isolation	21	17.3 (3.8)	-1.8 (1.2 ^a)	4.1(0.4 ^a)	.83
PhonemesNVR	Repetition	21	10.4 (3.7)	-.23	.02	.75
PhonemesNVR	Deletion	21	9.3 (3.6)	-.20	-.23	.73
PhonemesVR	Isolation	21	19.0 (2.4)	-2.5 (0.5 ^a)	9.8 (-1.2 ^a)	.73
PhonemesVR	Repetition	21	11.2 (3.2)	-.26	.03	.71
PhonemesVR	Deletion	21	8.2 (4.0)	-.19	-.76	.80
PseudowordsVR	Isolation	21	13.4 (3.7)	-.84	.79	.76
PseudowordsVR	Repetition	21	12.0 (4.3)	-.98	.40	.81
PseudowordsVR	Deletion	21	6.8 (4.9)	.22	-1.1	.87
Digit span		7	4.3 (0.8)	-.04	-.44	.79
Corsi block span		8	4.6 (0.9)	-.08	.06	
Matrices		33	8.3 (3.8)	1.1	1.6	
BAS regular word reading		43	26.4 (9.3)	-.81	-.36	.98
DTWRP regular word reading		30	13.8 (7.4)	.29	-1.0	.97 ^b
DTWRP nonword reading		30	18.7 (7.2)	-.60	-.79	.96 ^b

Note: BAS regular word reading = British Ability Scales 2 word reading test, total for regular words only. DTWRP = Diagnostic Test of Word Reading Processes.

^a Inverse transformation.

^b Published reliability.

distinctively apart from the others, therefore it was eliminated from analyses (Byrne, 2010).

2.7. Statistical analysis

Two structural equation models were built for which five fit statistics are reported. As χ^2 is known to be sensitive to sample size and is less reliable for large samples such as in the current study (Tabachnik & Fidell, 2007) we report four additional measures of fit; the Normed Fit Index (NFI), the Comparative Fit Index (CFI), the Incremental Index of Fit (IFI) and the root mean square error of approximation (RMSEA) and its 90% confidence interval. For the first three fit indices, values greater than .85 indicate a reasonable fit, greater than .90 represent a good fit, while values greater than .95 represent a very good fit. RMSEA values less than .06 indicate a good fit, values up to .08 indicate a reasonable fit, and values greater than .10 represent a poor fit (Tabachnik & Fidell, 2007). All models were built in AMOS 20.0 using maximum likelihood estimation (IBM, 2012).

3. Results

Table 1 shows descriptive and normality statistics for the exogenous and endogenous measures. Mean levels indicate that in each case, isolation was the easiest task, followed by repetition and then deletion. Matched sets of tasks were similar in difficulty with the exception of the TonesNVR tasks which were more difficult than the PhonemesNVR tasks.

Table 2 shows the correlations between measures. All correlations within sets of tasks were statistically significant ($p < .01$) indicating appropriateness for factor analysis (Tabachnik & Fidell, 2007). Correlations between the control variables and the Decoding factor were .28 (Digit Span), .08 (Corsi block span) and .35 (Matrices).

3.1. Confirmatory factor analyses: Creating latent variables to represent components

3.1.1. The Stimulus model

Latent variables (factors) were built that represented the three types of stimulus; Tones, Phonemes and Pseudowords.

Confirmatory factor analyses performed on these three factors, plus the three control variables (verbal short-term memory, visual-spatial short-term memory, and non-verbal reasoning) revealed a good fit to the data; $\chi^2(78) = 222.7$, $p < .001$, NFI = .905, IFI = .936, CFI = .935, RMSEA = .057, 90% CI = .048–.066.

3.1.2. The Response model

Latent variables were built that represented the two types of response; Non-verbal and Verbal. Confirmatory factor analyses performed on these two factors, plus the three control variables revealed a reasonable fit to the data; $\chi^2(83) = 484.8$, $p < .001$, NFI = .792, IFI = .821, CFI = .820, RMSEA = .092, 90% CI = .084–.100.

3.2. Structural equation modelling: Isolating task components in the prediction of decoding

3.2.1. The Stimulus model

An initial model with a direct link from just the Tone factor to Decoding provided an adequate fit to the data; $\chi^2(122) = 447.3$, $p < .001$, NFI = .897, IFI = .923, CFI = .923, RMSEA = .068, 90% CI = .062–.075. The link from Tones to Decoding was significant ($\beta = .56$, $p < .001$). When the Phonemes to Decoding link was added, it made a significant additional contribution ($\beta = .48$, $p < .001$), and the model fit significantly improved; $\Delta\chi^2(1) = 47.8$, $p < .001$. This model provided a good fit to the data; $\chi^2(121) = 399.5$, $p < .001$, NFI = .908, IFI = .934, CFI = .934, RMSEA = .064, 90% CI = .057–.071. Finally, when the Pseudoword to Decoding link was added, the model fit significantly improved; $\Delta\chi^2(1) = 68$, $p < .001$. This model provided a good fit to the data; $\chi^2(120) = 331.5$, $p < .001$, NFI = .924, IFI = .903, CFI = .950, RMSEA = .056, 90% CI = .049–.063. In the final model, the regression weight from Tones to Decoding was non-significant ($\beta = .06$, $p = .30$), while both the Phoneme ($\beta = .13$, $p < .05$) and Pseudoword to Decoding links were significant ($\beta = .53$, $p < .001$). See Fig. 2.

3.2.2. The Response model

A model with a direct link from just the non-verbal response factor to Decoding provided an adequate fit to the data; $\chi^2(127) = 652.6$, $p < .001$, NFI = .850, IFI = .876, CFI = .875, RMSEA = .085, 90% CI = .079–.092. This link was significant, ($\beta = .60$, $p < .001$). When the Verbal response to Decoding link

Table 2

Correlations between predictor, control, and outcome variables.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. TonesNVR isolation	–	.50	.47	.34	.35	.30	.29	.29	.19	.34	.18	.40	.25	.22	.33	.31	.32	.30
2. TonesNVR repetition		–	.57	.29	.34	.30	.19	.28	.21	.22	.05	.32	.31	.18	.33	.28	.29	.32
3. TonesNVR deletion			–	.29	.27	.31	.14	.22	.23	.19	.03	.28	.19	.19	.29	.21	.22	.24
4. PhonemesNVR isolation				–	.43	.40	.32	.25	.23	.18	.08	.30	.18	.21	.25	.26	.27	.26
5. PhonemesNVR repetition					–	.58	.32	.45	.31	.32	.17	.39	.29	.23	.34	.32	.33	.33
6. PhonemesNVR deletion						–	.30	.39	.44	.27	.11	.38	.27	.23	.31	.33	.31	.31
7. PhonemesVR isolation							–	.34	.27	.21	.20	.29	.17	.20	.20	.26	.25	.23
8. PhonemesVR repetition								–	.35	.27	.18	.35	.30	.15	.24	.35	.35	.32
9. PhonemesVR deletion									–	.19	.07	.33	.14	.23	.26	.29	.25	.26
10. PseudowordsVR isolation										–	.46	.49	.18	.17	.27	.40	.39	.32
11. PseudowordsVR repetition											–	.38	.11	.08	.13	.15	.14	.11
12. PseudowordsVR deletion												–	.30	.18	.39	.51	.53	.52
13. Digit span													–	.19	.24	.33	.35	.33
14. Corsi block span														–	.27	.20	.19	.17
15. Matrices															–	.41	.40	.42
16. BAS Regular word reading																–	.90	.78
17. DTWRP Regular word reading																	–	.85
18. DTWRP Nonword reading																		–

Note: Correlations are Bivariate (Pearson's r), $n = 570$, Critical $r = .10$ for significance at $p < .01$. BAS regular word reading = British Ability Scales 2 word reading score for regular words only. DTWRP = Diagnostic Test of Word Reading Processes.

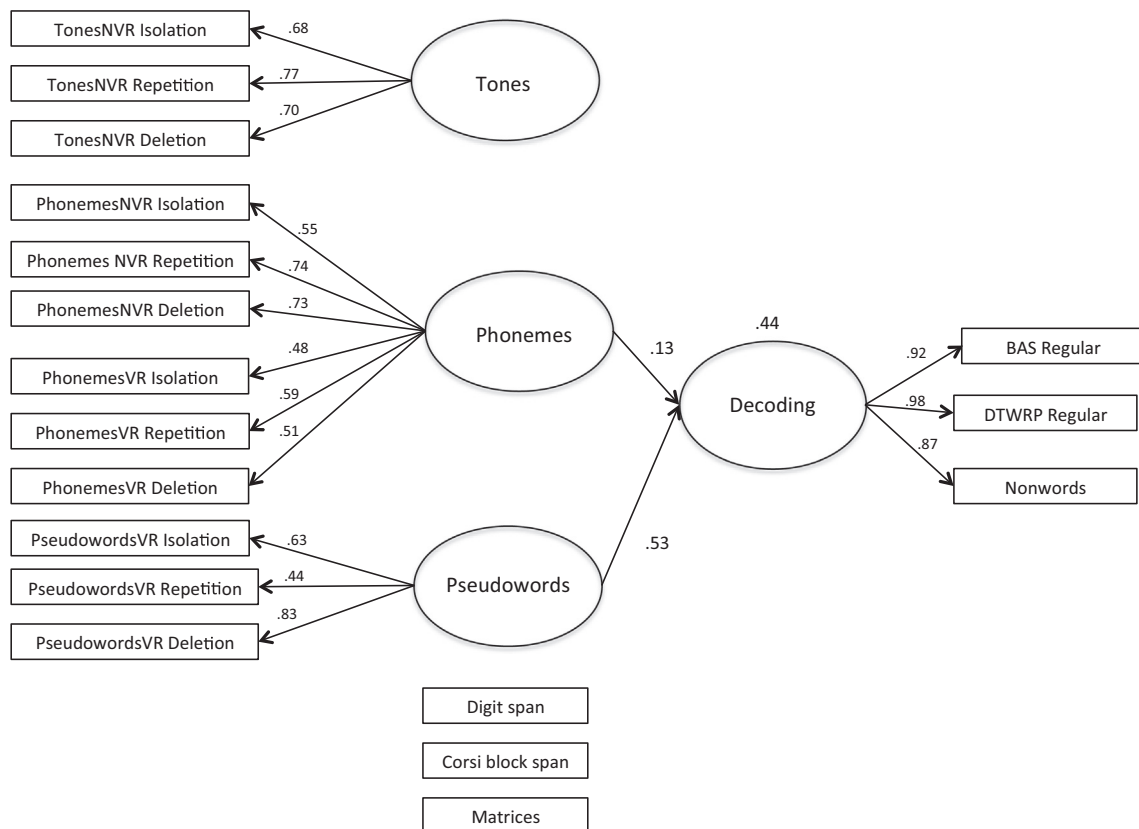


Fig. 2. The Stimulus model. Single-headed arrows leading away from the factors represent factor loadings (r). Single headed arrows leading to decoding represent regression weights (β). R^2 for the whole model is shown above Decoding. All correlations (covariances) between factors and between factors and control variables were modelled and were significant at the $p < .001$ level. For simplicity, covariances and error variances have been omitted from the diagram.

was added, the model fit significantly improved; $\Delta\chi^2(1) = 61.6$, $p < .001$. This model provided a reasonable fit to the data; $\chi^2(126) = 591.0$, $p < .001$, NFI = .864, IFI = .890, CFI = .889, RMSEA = .081, 90% CI = .074–.087. In the final model, the regression weight from Non-verbal response to Decoding was non-significant ($\beta = -.12$, $p = .22$), while the Verbal response to Decoding link was significant ($\beta = .79$, $p < .001$). See Fig. 3.

4. Discussion

Recent research has highlighted the importance of deconstructing complex tasks when investigating the skills that underlie reading (see Protopapas, 2014; Ramus & Ahissar, 2012 for discussions). Similar discussions have arisen in the context of other cognitive abilities, such as executive function and working memory (e.g.,

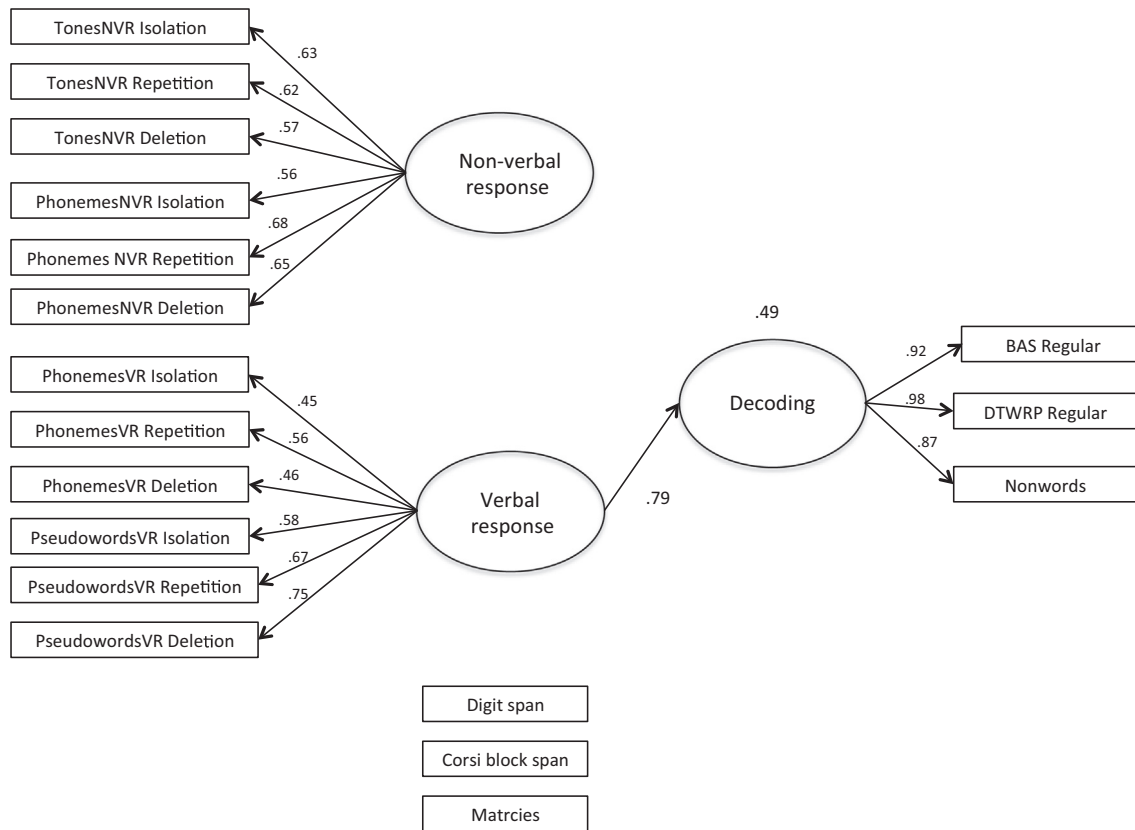


Fig. 3. The Response model. Single-headed arrows leading away from the factors represent factor loadings (r). Single headed arrows leading to Decoding represent regression weights (β). R^2 for the whole model is shown above Decoding. All correlations (covariances) between factors and between factors and control variables were modelled and were significant at the $p < .01$ level. For simplicity, covariances and error variances have been omitted from the diagram.

Conway et al., 2005; Hughes, 2011). In our study, we aimed to isolate specific components of phonological tasks and assess their contribution to the prediction of decoding skill. Two key advances were made; first, we outlined a method for separating the demands of complex tasks through combining a matched task design with structural equation modelling. Second, by isolating the separate contributions of each stimulus type and each response type, we were able to explore more precisely the nature of the relationship between phonological processes and reading.

4.1. The Stimulus model

According to phonological accounts, individual differences in processing and representing linguistic stimuli are fundamental in predicting reading attainments. The detailed, segmental nature of these representations is critical, and representation at the level of the phoneme is of particular importance (e.g., Melby-Lervag et al., 2012; Snowling, 2000; Snowling & Hulme, 1994). These components are embodied in classic phonological tasks, which require one to represent a speech code in short-term memory in a segmentally organised fashion, and in the case of isolation and deletion, explicitly break-up the code, and isolate a specific segment. Consistent with phonological accounts, we found that a factor representing simple linguistic stimuli (Phonemes) contributed unique variance over and above a factor representing closely matched comparison tasks without this feature (Tones). In addition, it was shown that a factor representing phonologically complex stimuli (Pseudowords) predicted unique variance in decoding over and above comparison tasks without this feature (Phonemes). These results held even when verbal short-term memory, along with other controls, was co-varied, showing that the effect of stimulus

was not driven by the increased memory load involved in processing linguistic stimuli (e.g., due to the stimuli being more complex; Ramus & Szenkovits, 2008). This finding highlights the importance of stimulus-specific features (the linguistic nature of the stimulus, and the phonological complexity of the stimulus) as important driving forces in the relationship between phonological tasks and reading.

It is of interest that the simple linguistic factor (Phonemes) predicted unique variance in decoding even once the link from the phonologically complex factor (Pseudowords) was included in the Stimulus model. This suggests that efficient access to both simple and complex phonological representations is important for decoding (see Ramus & Szenkovits, 2008). Since each factor was uniquely predictive, they must have challenged different processes. Whereas the pseudoword tasks demanded the ability to segment complex word-like stimuli, the phoneme tasks challenged access to discrete phonological representations, outside the context of a word. The ability to access discrete phonological representations may be particularly helpful in the early stages of reading, while a child is continuing to build up their store of orthographic knowledge. When a word is unfamiliar, and must be decoded at the level of the phoneme, access to representations of individual phonemes may be useful, over and above access to larger units or whole word representations (e.g. Ehri, 2005).

4.2. The Response model

Classic phonological tasks also require one to produce a detailed phonological output representation, access the relevant articulatory information and produce the motor movements required to pronounce the answer (e.g., see Ramus et al., 2010, for details of

the full information-processing model). Previous research has highlighted the importance of phonological output processes to reading. For example, Litt et al. (2013) found that a factor representing tasks with a verbal output fully explained the paired associate learning to reading relationship (matched tasks involving a written response were not significant predictors). In addition, other work has found evidence of a specific phonological output deficit in developmental dyslexia (Hulme & Snowling, 1992). Although there is an ongoing debate concerning the roles of the awareness and memory aspects of phonological processing (e.g., see Nithart et al., 2011), classic tasks clearly challenge both phonological memory and awareness as well as loading on phonological input and output processes. By comparing verbal and non-verbal response tasks that used common procedures and by partialing out the effect of verbal and visual-spatial STM, the Response model allowed us to specifically tap the role of phonological output processes while controlling for memory demands as closely as possible (e.g., using the mappings between stimulus and 'alien' in the non-verbal response tasks). In line with previous research, we found that the factor representing verbal response tasks predicted unique variance in decoding over and above the non-verbal response factor. This highlights the importance of producing a verbal response in the relationship between phonological tasks and reading.

Finally, the finding that the non-verbal response tasks in the Response model were *not* uniquely predictive implies that these tasks were 'neutral'. In other words, having the tasks in this format does not contribute anything additional in the context of our models. It does not mean that non-verbal skills on their own do not make a contribution to reading, as demonstrated by the significant effect of the non-verbal factor before the verbal factor was added to the model, and as evidenced by previous studies that have found non-verbal response phonological tasks to be good predictors of reading (e.g., Carroll, Snowling, Hulme, & Stevenson, 2003; Gayan & Olson, 2003).

4.3. Limitations

It is of note that we were not able to isolate the unique contribution of each component while *in competition* with all other components in the present study. Due to the nature of latent variable modelling (which was designed to model separate factors that show the best fit to the data), it was not possible to enter stimulus and response factors into the same model (as they would involve loadings onto the same tasks; each stimulus must have a response). Therefore, it was not possible to isolate unique contributions of stimulus and response at the same time. Instead, we isolated the importance of each type of stimulus in the Stimulus model and the separate importance of each type of response in the Response model. One possibility for future designs may be to include receptive tasks that do not require an explicit response in order to separate the two mechanisms.

Another point to make is that our task design was not fully factorial meaning that we were unable to investigate possible interactions between complex phonological stimuli and response mode, and between non-linguistic stimuli and response mode. Therefore we can only interpret the contribution of each component within the context in which it was investigated. Specifically, we found that phonologically complex stimuli (pseudowords) were uniquely important to decoding, over and above simple linguistic stimuli (phonemes) but only in the context of a verbal response. And we found that a verbal response was uniquely important to decoding, over and above matched non-verbal tasks, but only in the context of simple linguistic stimuli (phonemes). Future research would benefit from including additional conditions in their design.

Finally, although we successfully investigated three important components of phonological tasks in the present study, we did not isolate every task characteristic. Components that we did not explicitly manipulate include phonological memory, attention and speed. A similar approach could be applied to create matched tasks that vary along these dimensions in order to isolate their specific effect (e.g. by comparing shorter and longer pseudowords to manipulate phonological memory). These types of designs could be fruitfully combined with a latent variable modelling approach in future research.

5. Conclusions

In conclusion, we found that two stimulus-specific components of phonological tasks; (a) the simple linguistic nature of the stimulus, and (b) the phonological complexity of the stimulus (within the context of a verbal response) were uniquely predictive of early decoding outcomes, over and above comparison measures without these components. We also found that (c) the production of a verbal response (within the context of a linguistic stimulus), was uniquely predictive of early decoding skills, over and above comparison measures with a non-verbal response. Therefore, we suggest that all three components play a role in driving the relationship between phonological tasks and reading. The results support phonological theories of reading, highlighting the importance of both phonological input (creation and access to representations – the Stimulus model) and output (articulation – the Response model) processes to early reading skills. Finally, by demonstrating contributions from multiple stimulus and response factors, our findings provide fresh support for multivariate models of reading that go beyond classic single-factor explanations (e.g., Pennington, 2006 in the context of developmental disorders; Vellutino, Tunmer, Jaccard, & Chen, 2007 in the context of normal development).

This study demonstrated that using carefully matched tasks as indicators for latent variables provides a powerful approach to breaking down complex processes. Although we focussed on three of the most basic characteristics of phonological tasks in the current study, this approach could be used to address a number of outstanding questions about the nature of phonological tasks and their role in reading achievement (e.g., the roles of memory and speed). Of course the challenges of deconstructing complex tasks apply outside reading research, and we believe that the approach used here can serve as a valuable model in other domains, such as executive function and components of language.

References

- Audacity (2011). Audacity (Version 1.3.4-beta) [Computer program]. <<http://audacity.sourceforge.net/>> Retrieved 30.05.11.
- Banai, K., & Ahissar, M. (2006). Auditory processing deficits in dyslexia: Task or stimulus related? *Cerebral Cortex*, 16(12), 1718–1728. <http://dx.doi.org/10.1093/cercor/bhj107>.
- Bowey, J. A. (2007). Predicting individual differences in learning to read. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook*. Oxford: Blackwell Publishing.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS* (2nd ed.). Routledge.
- Carroll, J. M., Snowling, M. J., Hulme, C., & Stevenson, J. (2003). The development of phonological awareness in preschool children. *Developmental Psychology*, 39(5), 913–923.
- Conway, A. A., Kane, M., Bunting, M., Hambrick, D. Z., Wilhelm, O., & Engle, R. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769–786. <http://dx.doi.org/10.3758/bf03196772>.
- Cunningham, A. J., Witton, C., Talcott, J. B., Burgess, A. P., & Shapiro, L. R. (2011). The Aston literacy project. <<http://www.aston.ac.uk/alp/>> Retrieved 2.07.2015.
- de Jong, P. F. (1998). Working memory deficits of reading disabled children. *Journal of Experimental Child Psychology*, 70(2), 75–96. <http://dx.doi.org/10.1006/jecp.1998.2451>.
- De Lillo, C. (2004). Imposing structure on a Corsi-type task: Evidence for hierarchical organisation based on spatial proximity in serial-spatial memory. *Brain and Cognition*, 55(3), 415–426. <http://dx.doi.org/10.1016/j.bandc.2004.02.071>.

- Department for Education (2013). Schools, pupils and their characteristics. <<https://www.gov.uk/government/statistics/schools-pupils-and-their-characteristics-january-2013>> Retrieved 2.07.2015.
- Department for Education and Skills (2007). Letters and sounds: Principles and practice of high quality phonics.
- Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, 9(2), 167–188. http://dx.doi.org/10.1207/s1532799xssr0902_4.
- Elliot, C. D., Smith, P., & McCulloch, K. (1996). *British ability scales* (2nd ed.). Slough: National Foundation of Educational Research.
- Forum for Research in Literacy and Language (2012). Diagnostic test of word reading processes. Institute of Education: GL Assessment.
- Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, 2(2), 103–127. <http://dx.doi.org/10.1080/09658219408258940>.
- Gayan, J., & Olson, R. K. (2003). Genetic and environmental influences on individual differences in printed word recognition. *Journal of Experimental Child Psychology*, 84(2), 97–123. [http://dx.doi.org/10.1016/S0022-0965\(02\)00181-9](http://dx.doi.org/10.1016/S0022-0965(02)00181-9).
- Groth, K., Lachmann, T., Riecker, A., Muthmann, I., & Steinbrink, C. (2011). Developmental dyslexics show deficits in the processing of temporal auditory information in German vowel length discrimination. *Reading and Writing*, 24(3), 285–303. <http://dx.doi.org/10.1007/s11455-009-9213-7>.
- Hughes, C. (2011). Changes and challenges in 20 years of research into the development of executive functions. *Infant and Child Development*, 20(3), 251–271. <http://dx.doi.org/10.1002/icd.736>.
- Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological Science*, 23(6), 572–577. <http://dx.doi.org/10.1177/0956797611435921>.
- Hulme, C., & Snowling, M. J. (1992). Deficits in output phonology: An explanation of reading failure? *Cognitive Neuropsychology*, 9(1), 47–72. <http://dx.doi.org/10.1080/02643299208252052>.
- Hulslander, J., Talcott, J., Witton, C., DeFries, J., Pennington, B., Wadsworth, S., et al. (2004). Sensory processing, reading, IQ, and attention. *Journal of Experimental Child Psychology*, 88(3), 274–295. <http://dx.doi.org/10.1016/j.jecp.2004.03.006> [Article].
- IBM (2012). IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp. IBM.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133(2), 189–217. <http://dx.doi.org/10.1037/0096-3445.133.2.189>.
- Kessels, R. P., van Zandvoort, M. J., Postma, A., Kappelle, L. J., & de Haan, E. H. (2000). The Corsi Block-Tapping Task: Standardization and normative data. *Applied Neuropsychology*, 7(4), 252–258. http://dx.doi.org/10.1207/s15324826an0704_8.
- Kilminster, M., & Laird, E. (1978). Articulation development in children aged three to nine years. *Australian Journal of Human Communication Disorders*, 6(1), 23–30.
- Lervag, A., Braten, I., & Hulme, C. (2009). The cognitive and linguistic foundations of early reading development: A Norwegian latent variable longitudinal study. *Developmental Psychology*, 45(3), 764–781. <http://dx.doi.org/10.1037/A0014132>.
- Litt, R. A., de Jong, P. F., van Bergen, E., & Nation, K. (2013). Dissociating crossmodal and verbal demands in paired associate learning (PAL): What drives the PAL, reading relationship? *Journal of Experimental Child Psychology* (0). <http://dx.doi.org/10.1016/j.jecp.2012.11.012>.
- Majerus, S., Linden, M. V. d., Mulder, L., Meulemans, T., & Peters, F. d. r. (2004). Verbal short-term memory reflects the sublexical organization of the phonological language network: Evidence from an incidental phonotactic learning paradigm. *Journal of Memory and Language*, 51(2), 297–306. <http://dx.doi.org/10.1016/j.jml.2004.05.002>.
- Melby-Lervag, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, 138(2), 322–352. <http://dx.doi.org/10.1037/a0026744>.
- Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, 78(2), 622–646. <http://dx.doi.org/10.1111/j.1467-8624.2007.01018.x>.
- Moore, B. C. J. (2003). *An introduction to the psychology of hearing*. London: Academic Press.
- Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology*, 40(5), 665–681.
- Nation, K., & Hulme, C. (2011). Learning to read changes children's phonological skills: Evidence from a latent variable longitudinal study of reading and nonword repetition. *Developmental Science*, 14(4), 649–659. <http://dx.doi.org/10.1111/j.1467-7687.2010.01008.x>.
- Nithart, C., Demont, E., Metz-Lutz, M.-N., Majerus, S., Poncelet, M., & Leybaert, J. (2011). Early contribution of phonological awareness and later influence of phonological memory throughout reading acquisition. *Journal of Research in Reading*, 34(3), 346–363. <http://dx.doi.org/10.1111/j.1467-9817.2009.01427.x>.
- Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, 101(2), 385–413. <http://dx.doi.org/10.1016/j.cognition.2006.04.008> [Review].
- Protopapas, A. (2014). From temporal processing to developmental language disorders: Mind the gap. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1634). <http://dx.doi.org/10.1098/rstb.2013.0090>.
- Puolakanaho, A., Ahonen, T., Aro, M., Eklund, K., Leppanen, P. H. T., Poikkeus, A. M., et al. (2007). Very early phonological and language skills: Estimating individual risk of reading disability. *Journal of Child Psychology and Psychiatry*, 48(9), 923–931. <http://dx.doi.org/10.1111/j.1469-7610.2007.01763.x>.
- Ramus, F., & Ahissar, M. (2012). Developmental dyslexia: The difficulties of interpreting poor performance, and the importance of normal performance. *Cognitive Neuropsychology*, 29(1–2), 104–122. <http://dx.doi.org/10.1080/02643294.2012.677420>.
- Ramus, F., Marshall, C. R., Rosen, S., & van der Lely, H. K. J. (2013). Phonological deficits in specific language impairment and developmental dyslexia: Towards a multidimensional model. *Brain*, 136(2), 630–645. <http://dx.doi.org/10.1093/brain/aws356>.
- Ramus, F., Peperkamp, S., Christophe, A., Jacquemot, C., Kouider, S., & Dupoux, E. (2010). A psycholinguistic perspective on the acquisition of phonology. *Laboratory Phonology*, 10, 311–340.
- Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *Quarterly Journal of Experimental Psychology*, 61(1), 129–141. <http://dx.doi.org/10.1080/17470210701508822>.
- Rigdon, E. E. (1995). A necessary and sufficient identification rule for structural models estimated in practice. *Multivariate Behavioral Research*, 30(3), 359–383. http://dx.doi.org/10.1207/s15327906mbr3003_4.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburgh: Psychology Software Tools, Inc..
- Snow, C. E., & Juel, C. (2007). Teaching children to read: What do we know about how to do it? In M. J. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook*. Oxford: Blackwell publishing.
- Snowling, M. J. (2000). *Dyslexia* (2nd ed.). Oxford, UK: Blackwell.
- Snowling, M. J., Chiat, S., & Hulme, C. (1991). Words, nonwords, and phonological processes: Some comments on Gathercole, Willis, Emslie, and Baddeley. *Applied Psycholinguistics*, 12(3), 369–373. <http://dx.doi.org/10.1017/S0142716400009279>.
- Snowling, M. J., Stothard, S. E., Clarke, P., Bowyer-Crane, C., Harrington, A., & Truelove, E. (2009). *York assessment of reading for comprehension*. GL Assessment.
- Snowling, M. J., & Hulme, C. (1994). The development of phonological skills. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 346(1315), 21–27. <http://dx.doi.org/10.1098/rstb.1994.0124>.
- Sweigart, A. (2010). Invent your own computer games with python, 2nd ed.
- Tabachnik, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Pearson.
- Vandermosten, M., Boets, B., Luts, H., Poelmans, H., Wouters, J., & Ghesquière, P. (2011). Impairments in speech and nonspeech sound categorization in children with dyslexia are driven by temporal processing difficulties. *Research in Developmental Disabilities*, 32(2), 593–603. <http://dx.doi.org/10.1016/j.ridd.2010.12.015>.
- Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. S. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading*, 11(1), 3–32. http://dx.doi.org/10.1207/s1532799xssr1101_2.